DOCUMENT RESUME

ED 199 656 CS 005 947

AUTHOR Baumann, James F.
TITLE Research on Children's Reading Comprehension of Expository Prose: A Problem of Ecological Validity.
PUB DATE Jan 81
NOTE 18p.; Paper presented at the Annual Meeting of the Southwest Educational Research Association (Dallas, TX, January 1981).

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Classroom Environment; Classroom Research; Educational Research; Elementary Education; Generalization; Laboratory Experiments; Reading Comprehension; *Reading Research; Research Design; *Research Methodology; Research Needs; *Research Problems; *Research Utilization

ABSTRACT

There is some confusion in discriminating basic educational research (which explores the why's and how's of learning) from applied research (which seeks to test the implications of basic research results) because some investigations are conceived without a clear purpose. Thus, it is not uncommon for researchers to suggest implications that overstep their methodology. Research emphasizes internal validity, or highly controlled experiments, at the expense of more ecologically valid classroom testing. For example, implications from results of a reading comprehension study were applied to a classroom, but the original study involved adults in a laboratory setting, with an unusual comprehension measure. A replication of the experiment that corrected these deficiencies clearly proved that the classroom implications derived from the original study went far beyond the methodology. Credible implications can only come from studies designed with ecological validity. Researchers should replicate studies in a variety of contexts to ensure greater objectivity and inferences that are more applicable to the classroom. (HTH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED199656

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

Research on Children's Reading Comprehension of Expository Prose: A Problem of Ecological Validity

James F. Baumann

North Texas State University

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
James F. Baumann

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Paper presented at the Annual Meeting of the Southwest Educational Research Association

Dallas, January 1981

CS005947

Research on Children's Reading Comprehension of Expository Prose: A Problem of Ecological Validity

Educational researchers frequently classify studies by placing them in one of two categories: pure/basic investigations, or applied research efforts. The former category refers to investigations designed to identify or explain relationships between basic physiological, psychological, and social processes and learning. The latter category involves research designed to explore the efficacy of educational implications and/or applications that stem from basic research.

Intuitively, such a dichotomy is sensible. Some studies explore the why's and how's of human learning; other research seeks to determine how this knowledge can be applied to better educate people. In reality, however, identifying studies as either basic or applied is not so straightforward. This is true partly because no specific set of criteria clearly discriminate basic from applied research. In a sense, they reside more properly on a continuum as opposed to an either/or dichotomy.

But part of the confusion in discriminating basic from applied research is due to investigations that are conceived without a clear purpose--a purpose which should relate more to exploring underlying learning processes (basic educational research) or investigating pedagogical implications (applied educational research). And the purpose of a study dictates the kind of experimental method-

ology, which, in turn, delimits the kinds of conclusions and implications that are permissible. For example, if a researcher wished to determine how different kinds of typeface affect oral reading speed, a highly-controlled, laboratory study could be conducted to explore this question. If results indicated a strong relationship between Typeface A and fast oral reading, an important finding may have been uncovered. However, to suggest instructional implications--for example, to recommend that all basal readers should be printed in Typeface A--would be inappropriate. Further research involving children in schools with curricular materials printed in the variety of typefaces would be necessary before such statements would be justifiable.

Unfortunately, researchers frequently either have no clear purposes in mind for their studies, or they forget what they are. As a result, it is not uncommon for researchers to suggest implications or applications of results--usually of the educational variety--that go beyond the experimental methodology. That is, studies with limited generalizability, or ecological validity, are frequently used to make far-reaching educational statements.

It is the purpose of this paper to examine this problem and discuss the concept of ecological validity. An example within the research area of children's prose comprehension will be used to illustrate the problem. Suggestions for the design of more generalizable educational research will also be discussed.

Ecological Validity: What Is It?

Campbell and Stanley (1966) refer to two different forms of validity: internal and external validity. Internal validity involves the scientific rigor of an experiment: "Did in fact the experimental treatments make a difference in this specific experimental instance?" (Campbell & Stanley, 1966, p. 5). In other words, with how much confidence can a researcher attribute any treatment effects to the independent variable as opposed to extraneous, nuisance variables? External validity refers to the generalizability of results: "To what populations, settings, treatment variables, and measurement variables can this effect be generalized?" (Campbell & Stanley, 1966, p. 5).

Bronfenbrenner (1976) has taken the notion of external validity and articulated it further. He maintains that if educational research is to have any true impact, it must be taken out of the laboratory and conducted in real-life educational settings. In addition, research must take into account the ecology of the educational environment: the social, political, economic, linguistic, etc. context within which it is conducted. In so doing, Bronfenbrenner maintains, one is able to conduct a more natural experiment which will permit greater generalizability or ecological validity.

Historically, researchers have considered the constructs of internal and external validity as being inversely proportional:

the more there is of one, the less there is of the other. Since internal validity has been considered prerequisite for any experiment (the sine qua non as Campbell and Stanley state), the emphasis has been on experimental control rather than on generalizability of results. Journal editors have generally shared this bias and rewarded with publication studies possessing high levels of internal validity, the implicit assumption being that more externally valid research is "dirty" and results should be viewed skeptically. Consequently, there has been little motivation to conduct more ecologically valid experiments, and what one typically finds in educational research journals are studies possessing high internal control accompanied by discussion sections that contain educational implications/applications that seriously overstep the limits of the studies. This practice obviously is dangerous and oftentimes leads to unjustified or erroneous educational implications. The following section presents one such example.

An Example: Children's Reading Comprehension of Expository Prose

A popular and fruitful line of discourse processing research within the last decade has investigated the impact text structure variables has on reading comprehension. Researchers such as Walter Kintsch and Bonnie Meyer have argued that the manner in

which ideas are organized in expository texts influences the quantity and quality of those ideas remembered after reading a passage. Based on recent advancements in linguistics (e.g., Fillmore, 1968; Grimes, 1975), a number of formal systems capable of describing the content and relationships between ideas in prose have been developed (e.g., Frederiksen, 1975; Kintsch, 1974; Meyer, 1975). Although these prose analysis techniques differ in the specific manner in which ideas are classified and related, all systems result in a hierarchical description of text; that is, some ideas are central and thus emerge at superordinate positions in the passage hierarchy, and other ideas are more elucidative and thus reside at more subordinate levels. Furthermore, it has been hypothesized that those ideas most central to a text--those residing in the upper eshelons of the passage hierarchy--are most memorable.

Research by Kintsch and his associates (Kintsch & Keenan, 1973; Kintsch, Kozminsky, Streby, McKoon, & Keenan, 1975), Meyer and her colleagues (Meyer, 1975, 1977a, 1977b; Meyer & McConkie, 1973; Meyer, Rice, Knight, & Jessen, Note 1), and by others (Britton, Meyer, Simpson, Holdredge, & Curry, 1979; Duchastel, 1979) has supported this hierarchical hypothesis: Adult, fluent readers free recall significantly gre.ter amounts of High-level ideas compared to Low-level ideas; that is, height of information in the passage hierarchy ("content structure" as Meyer, 1975,

labels it) is directly related to recall. Meyer (1977a, study 3) also reported that children are sensitive to ideational prominence: After listening to an expository text, sixth graders correctly answered significantly greater numbers of questions that tapped information in the top-half of the text hierarchy than similar questions probing low-level ideas.

Based on the apparent robustness of this phenomenon (commonly referred to as the "levels effect"), the primary researcher in this area, Bonnie Meyer, offered a number of educational implications. Meyer (1975) suggested that the use of content structure tree diagrams would help students understand how information is linked together in a passage and help identify what concepts the author deemed most important. Meyer (1977a) suggested that the use of text hierarchies would supply comprehension test authors a rationale for writing various kinds and numbers of test questions. And more recently, Meyer (Note 2) specified an extensive set of educational recommendations concerning the writing and selection of content area textbooks and for instruction to improve reading comprehension.

Although Meyer is to be commended for her attempts to apply research findings to real-life educational settings, there are several conditions that limited the generalizability of the extant body of research cited above. This is especially true if the educational environment of interest is reading comprehension

instruction in the elementary school.

Specifically, four sets of factors have limited the external validity of the majority of research examining the levels hypothesis. First, virtually all of the existing studies involved adult, fluent readers. Consequently, inferences to reading educational programs for children are not backed by a strong empirical data base. Secondly, most of the extant research was conducted in highly-controlled laboratory settings, not naturalistic educational environments (i.e., schools). In addition, some studies (e.g., Meyer, 1977a, study 3; Meyer & McConkie, 1973) assessed subjects' listening comprehension rather than their reading comprehension. Thus, generalizing to children's reading comprehension behavior in schools becomes tenuous indeed. Thirdly, the experimental passages used in this line of research have been both experimenter contrived and limited in number making it an unjustifiable extrapolation to assume that children will exhibit the same comprehension behaviors when reading the wide variety of natural texts encountered in schools. And finally, the primary dependent variable used to measure comprehension in almost all of the studies was a free recall test (relate orally or in writing everything one can remember after reading a text): a task not commonly used to assess children's reading comprehension in elementary schools.

Because of the lack of ecological validity in this body of research, this researcher (Baumann, 1980) conducted a modified replication of Meyer's (1975) seminal study. Limitations in the four areas outlined above were addressed in the following manner: (1) Children (third and sixth graders) rather than adults participated in the study. (2) The study was conducted in schools in classrooms rather than in a laboratory environment. (3) The experimental passages consisted of unaltered natural texts that were randomly selected from pools of passages taken from the school environment (social studies textbooks). (4) In addition to a free recall test, children's reading comprehension was assessed by more typical classroom devices: paper and pencil question/answer tests.

Results of the study demonstrated little support for the levels hypothesis as a suitable descriptor of children's reading comprehension of expository prose. In only one instance--third graders responses to the multiple choice test--were the results consistent with the prior research; that is, greater comprehension of High-level, gist information. Across all other dependent measures (i.e., free recall test, short answer test, multiple choice test) at both grade levels no significant difference indicating the superior comprehension of High-level information was detected. In fact, there were several significant differences in which Low-level, detailed ideas were comprehended

better than High-level information. Thus, it was concluded that for the social studies texts used in the study, third and sixth grade children were unable to spontaneously comprehend central thoughts in expository prose with any great facility.

Conclusions

What is the significance of the above example? First, it demonstrates that even though an empirical finding has been consistently documented with a certain population of subjects under very specific conditions, one should not assume that the effect will generalize to other populations under different conditions. That is, just because adults were able to free recall greater amounts of High-level ideas after reading a few selected texts, it does not follow that children reading natural prose under normal school conditions will exhibit the same behavior in response to several dependent measures. The implicit caveat to researchers and readers of research which can be gleaned from this example is that educational recommendations to subjects, settings, materials, and dependent variables different from those employed in a study are at the very least tenuous indeed. Hence, researchers who desire to make credible educational implications, need to design studies with at least minimal levels of ecological validity.

One must remember, however, that ecological validity exists on a continuum and is not an all-or-nothing proposition; and, that to a certain extent, one does sacrifice internal validity in the process of enhancing generalizability. In the replication study described above (Baumann, 1980), for example, several concessions in the area of experimental design were necessary in order to achieve greater generalizability. One concession involved the selection of experimental passages. In order to be able to generalize to the realm of natural texts, passages were randomly selected. Even though the passages within the selection pools were restricted in several ways (e.g., readability, length), several passage-specific variables went uncontrolled and were treated as random variables. Thus, it could be argued that the failure to detect a levels effect in this study was due to the interaction or "cluttering" of a whole series of uncontrolled factors. In fact, a post hoc analysis of several of these variables (e.g., highlighting or signaling of ideas, the relationship among ideas or rhetorical structure) revealed that indeed the supposedly aberrant results could be explained by several of these factors.

But, in a sense, this analysis only strengthed the findings of the study in light of the purpose of the investigation which was to determine if children reading typical expository materials

will automatically comprehend a preponderance of High-level ideas. Since the results indicated that children are not inherently skillful at this task after reading natural texts--that is, passages in which these factors are not formally controlled--the educational suggestions that were beginning to stem from this assumption were all the more suspect.

It should also be noted that there may be instances in which the benefits in terms of added generalizability accrued from a certain methodological condition are outweighed by the concession in terms of internal validity. In the replication study, for example, testing was conducted by the experimenter as opposed to regular classroom teachers. In addition, testing was either done individually (for the free recall condition, by necessity) or in small groups. The combination of experimenter intervention and individual/small group testing may have affected the students' performance, for these are not conditions under which children's reading comprehension is normally assessed in schools. When researchers make decisions like this, they should be listed as limitations, and implications should be appropriately tempered when necessary. In fact, in spite of the level of ecological validity, educational implications should be stated with a good deal of "probably's" and "most likely's."

Perhaps a second "lesson" can be learned from the above example: Researchers should more frequently replicate studies,

and replicate them in a variety of contexts. No single one-shot experiment or series of similar studies should serve to prove or disprove any theory or justify any educational practice. Yet, unfortunately this frequently occurs, for the extrinsic rewards for replication are virtually nonexistent. Rather, researchers are encouraged to move on in a line of research as opposed to evaluating their findings again within different contexts. Pearson (1978) addressed this very issue of the failure to replicate research:

> Somehow the data get lost, but the conclusions live on, regenerating themselves. . . . Replication ought to be encouraged; however, our present system of refereed publication systematically discourages it (e.g., "this article adds nothing new to the literature"). (pp. 355-56)

In conclusion, it should be noted that there is a need for highly-controlled, laboratory studies, for from these efforts come the seeds for new and innovative educational practices. The danger enters, however, when researchers fail to take their studies out of the laboratory and into the classroom in order to assess the efficacy of their findings in real-life educational settings. Only after this most critical step are suggestions for educational practice truly justified. As Bronfenbrenner (1976) so eloquently states, in terms of educational applicability, the

need for natural or ecologically valid experimentation is present and great:

> There are many instances in which a design exploiting an experiment of nature provides a more critical contrast, insures greater objectivity, permits more precise and theoretically significant inferences--in short, is more elegant and constitutes "harder" science than the best possible contrived experiment addressed to the same research question. (p. 5)

Reference Notes

1. Meyer, B. J. F., Rice, G. E., Knight, C. C., & Jessen, J. L. Differences in the type of information remembered from prose by young, middle, and old adults. Arizona State University, Department of Educational Psychology, Prose Learning Series, Research Report No. 5, Summer 1979.

2. Meyer, B. J. F. Structure of prose: Implications for teachers of reading. Paper presented at the twenty-fourth annual convention of the International Reading Association, Atlanta, April 1979.

References

Baumann, J. F. Effects of height of information in the content structure and measurement mode on third and sixth graders' comprehension of expository discourse. Unpublished doctoral dissertation, University of Wisconsin-Madison, 1980.

Britton, B. K., Meyer, B. J. F., Simpson, R., Holdridge, T. S., & Curry, C. Effects of the organization of text on memory: Tests of two implications of the selective attention hypothesis. Journal of Experimental Psychology: Human Learning and Memory, 1979, 5(4), 496-506.

Bronfenbrenner, U. The experimental ecology of education. Educational Researcher, 1976, 5(9), 5-15.

Campbell, D. T., & Stanley, J. C. Experimental and quasi-experimental designs for research. Chicago: Rand McNally, 1966.

Duchastel, P. Learning objectives and the organization of prose. Journal of Educational Psychology, 1979, 71, 100-106.

Fillmore, C. J. The case for case. In E. Bach & R. T. Harms (Eds.), Universals in linguistic theory. New York: Holt, Rinehart and Winston, 1968.

Frederiksen, C. H. Representing logical and semantic structure of knowledge acquired from discourse. Cognitive Psychology, 1975, 7, 371-458.

Grimes, J. E. The thread of discourse. The Hague: Mouton, 1975.

Kintsch, W. The representation of meaning in memory. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1974.

Kintsch, W., & Keenan, J. M. Reading rate and retention as a function of the number of propositions in the base structure of sentences. Cognitive Psychology, 1973, 5, 257-274.

Kintsch, W., Kozminsky, E., Streby, W. J., McKoon, G., & Keenan, J. M. Comprehension and recall of text as a function of content variables. Journal of Verbal Learning and Verbal Behavior, 1975, 14, 196-214.

Meyer, B. J. F. The organization of prose and its effects on memory. Amsterdam: North-Holland, 1975.

Meyer, B. J. F. The structure of prose: Effects on learning and memory and implications for educational practice. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), Schooling and the acquisition of knowledge. Hillsdale, N.J.: Lawrence Erlbaum Associates, 1977. (a)

Meyer, B. J. F. What is remembered from prose: A function of passage structure. In R. O. Freedle (Ed.), Discourse production and comprehension. Norwood, N.J.: Ablex Publ. Co., 1977. (b)

Meyer, B. J. F., & McConkie, G. W. What is recalled after hearing a passage? Journal of Educational Psychology, 1973, 65, 109-117.

Pearson, P. D. On bridging gaps and spanning chasms. Curriculum Inquiry, 1978, 8(4), 353-363.